Rare events simulation: classical engineering methods and current trends using meta-models

B. Sudret

Chair of Risk, Safety & Uncertainty Quantification
Dpt of Civil, Environmental and Geomatic Engineering, ETH Zurich, Switzerland

October 26th, 2012

Some common engineering structures

Cattenom nuclear power plant (France)

Cormet de Roselend dam (France)

Military satellite

Airbus A380

Bladed disk

Computational models

- Modern engineering has to address problems of increasing complexity in various fields including infrastructures (civil engineering), energy (civil/mechanical engineering), aeronautics, defense, etc.

- Complex systems are designed using computational models that are based on:
  - a mathematical description of the physics (*e.g.* mechanics, acoustics, heat transfer, electromagnetism, etc.)
  - numerical algorithms that solve the resulting set of (usually partial differential) equations: finite element-, finite difference-, finite volume- methods, boundary element methods)

Computational models

Simulation models are calibrated and validated through comparison with lab experiments and *in situ* / full scale measurements. Once they are validated, these models may be run with different sets of input parameters in order to:

- explore the design space at low cost
- optimize the system w.r.t to cost criteria
- assess the robustness of the system w.r.t. uncertainties

Sources of uncertainty

- Differences between the designed and the real system in terms of material/physical properties and dimensions (tolerancing)
- Unforecast exposures: exceptional service loads, natural hazards (earthquakes, floods), climate loads (hurricanes, snow storms, etc.).

Global framework for managing uncertainties

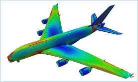
Step B
Quantification of sources of uncertainty

Step A
Model(s) of the system
Assessment criteria

Step C
Uncertainty propagation

Random variables

Computational model



Moments
Probability of failure
Response PDF

Step C'
Sensitivity analysis

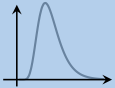
Global framework for managing uncertainties

Step B
Quantification of
sources of uncertainty

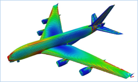
Step A
Model(s) of the system
Assessment criteria

Step C
Uncertainty propagation

Random variables



Computational model



Moments
Probability of failure
Response PDF

Step C'
Sensitivity analysis

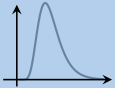
Global framework for managing uncertainties

Step B
Quantification of sources of uncertainty

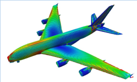
Step A
Model(s) of the system
Assessment criteria

Step C
Uncertainty propagation

Random variables



Computational model



Moments
Probability of failure
Response PDF

Step C'
Sensitivity analysis

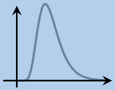
Global framework for managing uncertainties

Step B
Quantification of sources of uncertainty

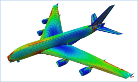
Step A
Model(s) of the system
Assessment criteria

Step C
Uncertainty propagation

Random variables



Computational model



Moments
Probability of failure
Response PDF

Step C'
Sensitivity analysis

Step A: computational models (civil & mechanical engineering)

Vector of input parameters $\boldsymbol{x} \in \mathbb{R}^M$

Computational model $\mathcal{M}$

Model response $\boldsymbol{y} = \mathcal{M}(\boldsymbol{x}) \in \mathbb{R}^N$

- geometry
- material properties
- loading

- analytical formula
- finite element model
- etc.

- displacements
- strains, stresses
- temperature, etc.

Step A: computational models (civil & mechanical engineering)

Vector of input parameters $\boldsymbol{x} \in \mathbb{R}^M$ — Computational model $\mathcal{M}$ — Model response $\boldsymbol{y} = \mathcal{M}(\boldsymbol{x}) \in \mathbb{R}^N$

- geometry
- material properties
- loading

- analytical formula
- finite element model
- etc.

- displacements
- strains, stresses
- temperature, etc.

Step B: probabilistic models of input parameters

No data exist

- expert judgment for selecting the input PDF's of $\boldsymbol{X}$
- literature, data bases (*e.g.* on material properties)
- maximum entropy principle

Input data exist

- classical statistical inference
- Bayesian statistics when data is scarce but there is some prior information

Data on output quantities

- inverse probabilistic methods and Bayesian updating techniques

Step B: probabilistic models of input parameters

No data exist

- expert judgment for selecting the input PDF's of $\boldsymbol{X}$
- literature, data bases (*e.g.* on material properties)
- maximum entropy principle

Input data exist

- classical statistical inference
- Bayesian statistics when data is scarce but there is some prior information

Data on output quantities

- inverse probabilistic methods and Bayesian updating techniques

Step B: probabilistic models of input parameters

No data exist

- expert judgment for selecting the input PDF's of $\boldsymbol{X}$
- literature, data bases (*e.g.* on material properties)
- maximum entropy principle

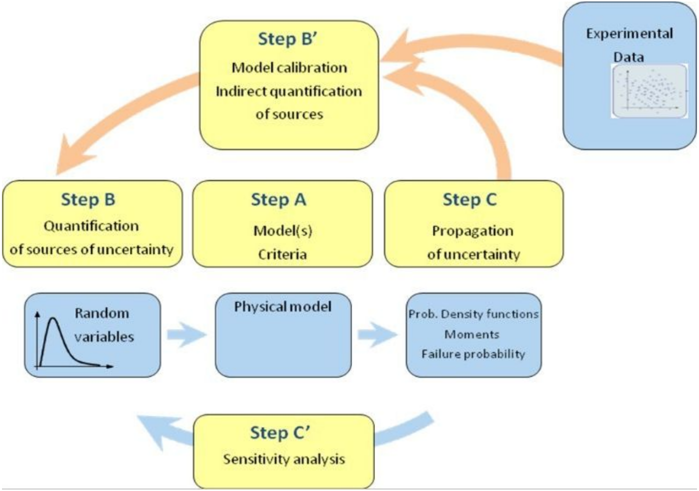
Input data exist

- classical statistical inference
- Bayesian statistics when data is scarce but there is some prior information

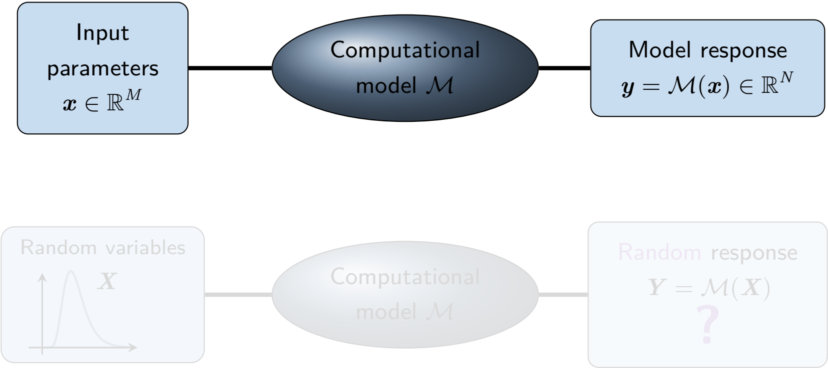
Data on output quantities

- inverse probabilistic methods and Bayesian updating techniques

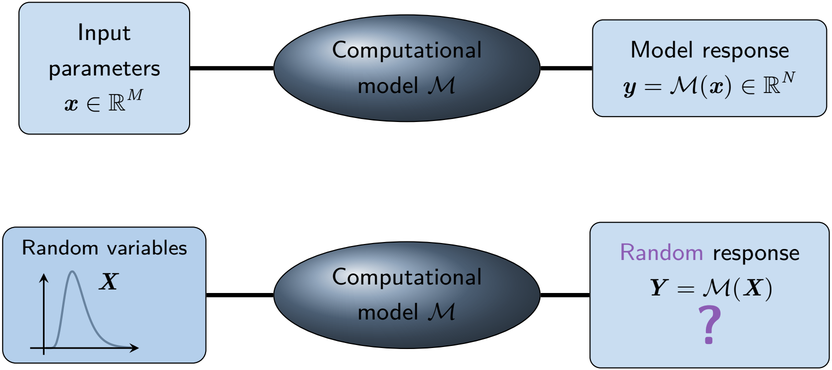
Step B: stochastic inverse problems



Step C: principles of uncertainty propagation



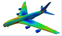
Step C: principles of uncertainty propagation



Step C: uncertainty propagation methods

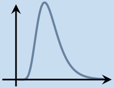
Computational model

Step A



Probabilistic model

Step B



Probabilistic-computational model

Mean/std. deviation

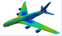
Rare event simulation

Response PDF

Step C: uncertainty propagation methods

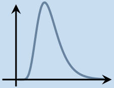
Computational model

Step A



Probabilistic model

Step B



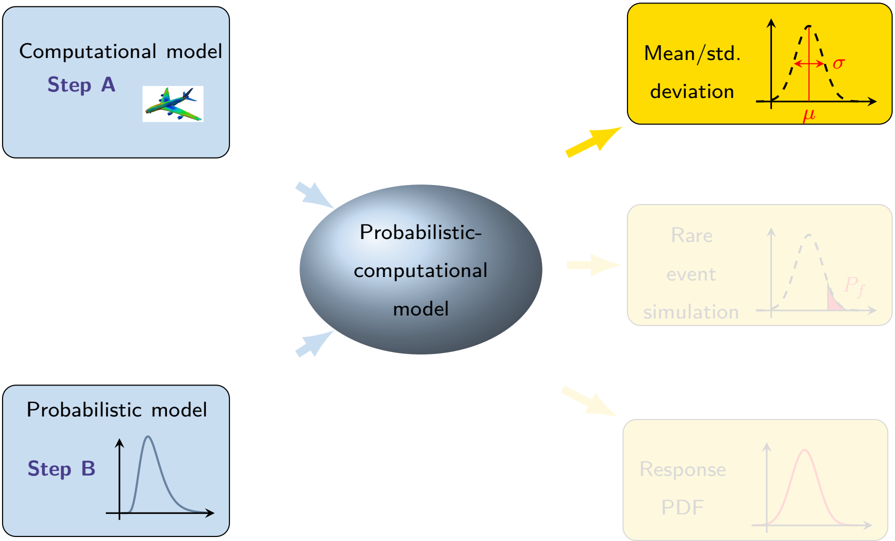
Probabilistic-computational model

Mean/std. deviation

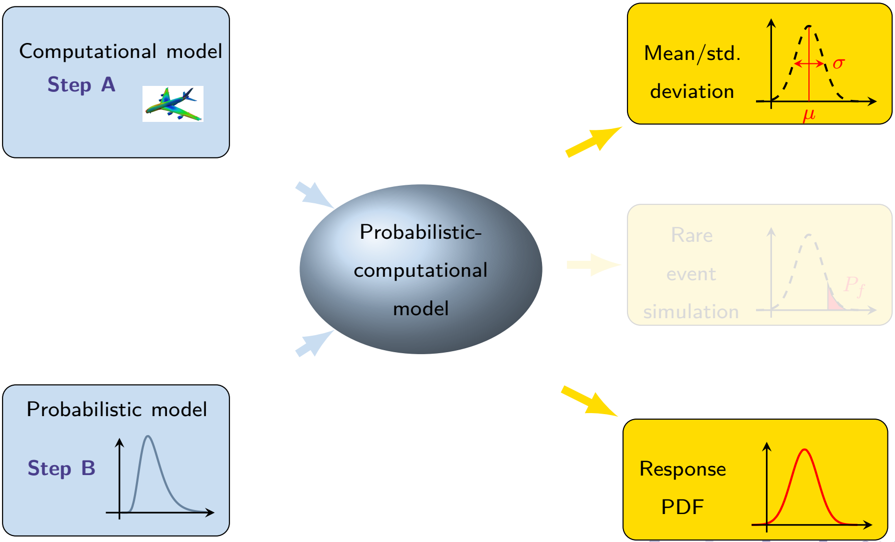
Rare event simulation

Response PDF

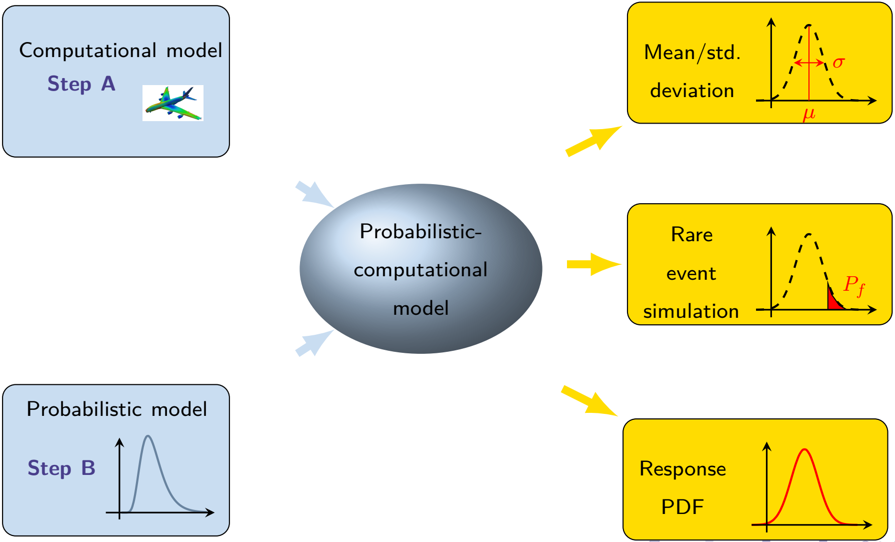
Step C: uncertainty propagation methods



Step C: uncertainty propagation methods



Step C: uncertainty propagation methods



Limit state function

- For the assessment of the system's performance, failure criteria are defined, *e.g.* :

$$\text{Failure} \quad \Leftrightarrow \quad q = \mathcal{M}(\boldsymbol{X}) \geq q_{adm}$$

Examples:

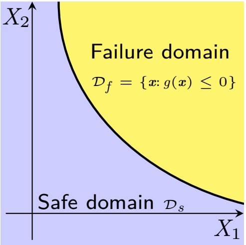
  - admissible stress / displacements in civil engineering
  - max. temperature in heat transfer problems
  - crack propagation criterion in fracture mechanics

- The failure criterion is cast as a limit state function (performance function) $g: \ \boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}} \mapsto \mathbb{R}$ such that:

$g(\boldsymbol{x}, \mathcal{M}(\boldsymbol{x})) \leq 0$ Failure domain $\mathcal{D}_f$

$g(\boldsymbol{x}, \mathcal{M}(\boldsymbol{x})) > 0$ Safety domain $\mathcal{D}_s$

$g(\boldsymbol{x}, \mathcal{M}(\boldsymbol{x})) = 0$ Limit state surface

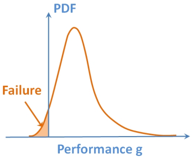


Probability of failure

The probability of failure is defined by:

$$P_f = \mathbb{P}\left(\{\boldsymbol{X} \in D_f\}\right) = \mathbb{P}\left(g\left(\boldsymbol{X}, \mathcal{M}(\boldsymbol{X})\right) \leq 0\right)$$

$$P_f = \int_{\mathcal{D}_f} f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x}$$



Features

- P_f is defined as a multidimensional integral, whose dimension is equal to the number of basic input variables $M = \dim \boldsymbol{X}$.
- The domain of integration is not known explicitly: it is defined by a condition related to the sign of the limit state function, which depends itself on the basic variables through a (potentially complex) mechanical model:

$$\mathcal{D}_f = \{\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{X}} : \quad g(\boldsymbol{x}, \mathcal{M}(\boldsymbol{x})) \leq 0\}$$

- Failures are (usually) rare events: the probability of interest typically ranges from 10^{-2} to 10^{-8}.

Outline

Monte Carlo simulation
Basic equations

- Let us introduce the indicator function of the failure domain:

$$\mathbf{1}_{\mathcal{D}_f}(\boldsymbol{x}) = \begin{cases} 1 & \text{if } g\left(\boldsymbol{x}, \mathcal{M}\left(\boldsymbol{x}\right)\right) \leq 0 \\ 0 & \text{otherwise} \end{cases}$$

- The probability of failure reads:

$$\begin{aligned} P_f &= \int_{D_f = \{\boldsymbol{x}\,:\, g(\boldsymbol{x}, \mathcal{M}(\boldsymbol{x})) \leq 0\}} f_{\boldsymbol{X}}(\boldsymbol{x})\; d\boldsymbol{x} \\ &= \int_{\mathbb{R}^M} \mathbf{1}_{\mathcal{D}_f}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\; d\boldsymbol{x} = \mathbb{E}\left[\mathbf{1}_{\mathcal{D}_f}(\boldsymbol{X})\right] \end{aligned}$$

- The following estimator is used:

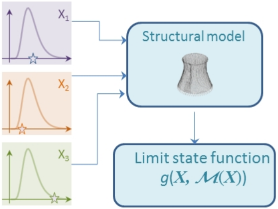
$$\hat{P}_f = \frac{1}{N} \sum_{i=1}^{N} 1_{\mathcal{D}_f}\left(\boldsymbol{X}_i\right) \qquad \boldsymbol{X}_i \;:\; \text{i.i.d copies of } \boldsymbol{X}$$

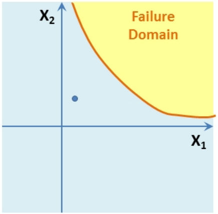
Estimator of the probability of failure P_f

- A sample set of input parameters $\mathcal{X} = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_N\}$, is drawn. For each sample the model response is computed and the limit state function $g\left(\boldsymbol{x}_i, \mathcal{M}\left(\boldsymbol{x}_i\right)\right)$ is evaluated.

Estimator of the probability of failure P_f

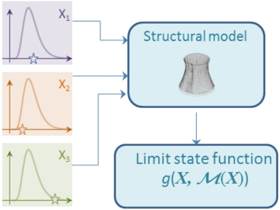
- A sample set of input parameters $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$, is drawn. For each sample the model response is computed and the limit state function $g(\boldsymbol{x}_i, \mathcal{M}(\boldsymbol{x}_i))$ is evaluated.

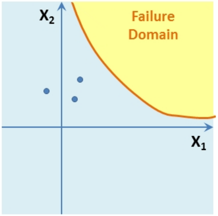




Estimator of the probability of failure P_f

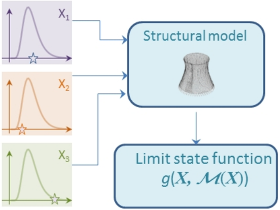
- A sample set of input parameters $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$, is drawn. For each sample the model response is computed and the limit state function $g(\boldsymbol{x}_i, \mathcal{M}(\boldsymbol{x}_i))$ is evaluated.

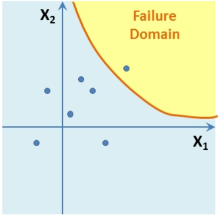




Estimator of the probability of failure P_f

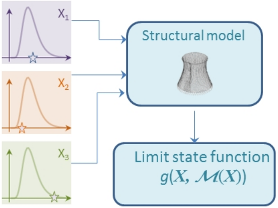
- A sample set of input parameters $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$, is drawn. For each sample the model response is computed and the limit state function $g(\boldsymbol{x}_i, \mathcal{M}(\boldsymbol{x}_i))$ is evaluated.

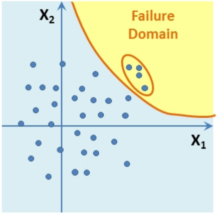




Estimator of the probability of failure P_f

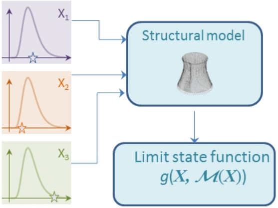
- A sample set of input parameters $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$, is drawn. For each sample the model response is computed and the limit state function $g\left(\boldsymbol{x}_i, \mathcal{M}\left(\boldsymbol{x}_i\right)\right)$ is evaluated.

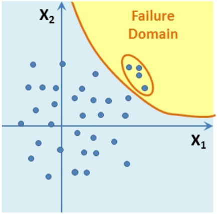




Estimator of the probability of failure P_f

- A sample set of input parameters $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$, is drawn. For each sample the model response is computed and the limit state function $g(\boldsymbol{x}_i, \mathcal{M}(\boldsymbol{x}_i))$ is evaluated.





- The number of negative values of the g-function, say N_f is stored. and P_f is estimated by:

$$P_f = \frac{N_f}{N}$$

Estimator of the probability of failure P_f

- The estimator $\hat{P}_f$ is a sum of Bernoulli variables: it has a binomial distribution with mean value $\mathbb{E}\left[\hat{P}_f\right] = P_f$ (unbiasedness) and variance $\text{Var}\left[\hat{P}_f\right] = \frac{1}{N} P_f\,(1 - P_f)$.
- Its coefficient of variation reduces to $CV \approx 1/\sqrt{N\,P_f}$ for rare events.

The convergence rate of Monte Carlo simulation is $\propto 1/\sqrt{N}$

Minimal size of the sample set

Suppose the probability of failure under consideration is of magnitude $P_f = 10^{-k}$ and an accuracy of 5% is aimed at.

$$CV_{P_f} = \frac{1}{\sqrt{N\,P_f}}$$

$$CV_{P_f} \leq 5\% \Longrightarrow N \geq 4.10^{k+2}$$

P_f	N_{min}
10^{-2}	40,000
10^{-3}	400,000
10^{-4}	4,000,000
10^{-6}	400,000,000

Outline

Introduction

Principle

The First Order Reliability Method (FORM) aims at approximating the integral which defines the probability of failure. It relies upon three steps:

- an iso-probabilistic transform of the input random vector $\boldsymbol{X}$ into a standard normal vector $\boldsymbol{U}$
- the search for the design point in this space
- the linearization of the limit state surface at the design point and the computation of the approximated failure probability

Step 1: iso-probabilistic transform

Principle

- The input random vector $\boldsymbol{X}$ is transformed into a standard normal random vector $\boldsymbol{U}$. Let us denote by $\mathcal{T}$ the iso-probabilistic transform:

$$\boldsymbol{X} \sim f_{\boldsymbol{X}} \qquad \boldsymbol{X} = \mathcal{T}(\boldsymbol{U}) \quad \text{where } \boldsymbol{U} \sim \mathcal{N}(\mathbf{0}, \mathbf{I_M})$$

- This reduces to a mapping of the integral from the physical space (that of $\boldsymbol{X}$) to the standard normal space (that of $\boldsymbol{U}$):

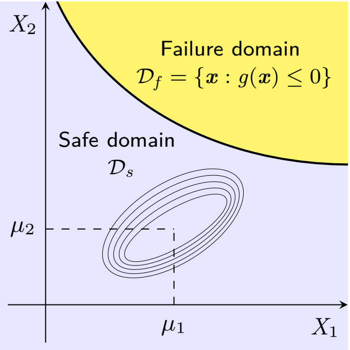
$$P_f = \int_{D_f = \left\{\boldsymbol{u} \in \mathbb{R}^M \,:\, g(\mathcal{T}(\boldsymbol{u})) \leq 0\right\}} \varphi_M(\boldsymbol{u}) \, d\boldsymbol{u}$$

where the standard normal PDF reads:

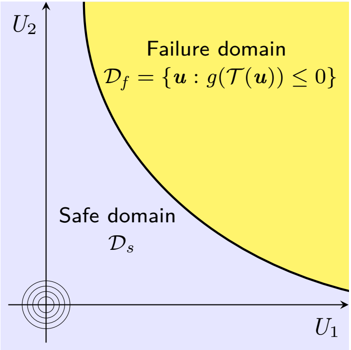
$$\varphi_M(\boldsymbol{u}) = (2\pi)^{-M/2} \exp\left[-\frac{1}{2}(u_1^2 + \cdots + u_M^2)\right]$$

Step 1: iso-probabilistic transform

Illustration



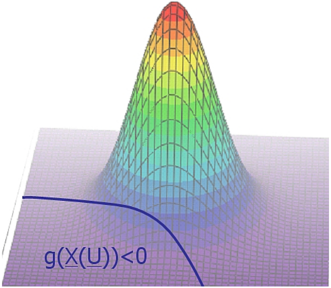
Physical space

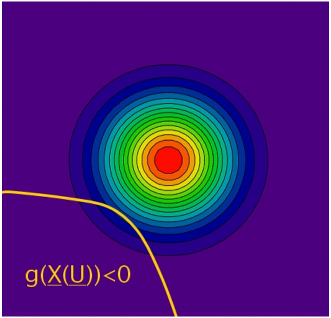


Standard normal space

Step 1: iso-probabilistic transform

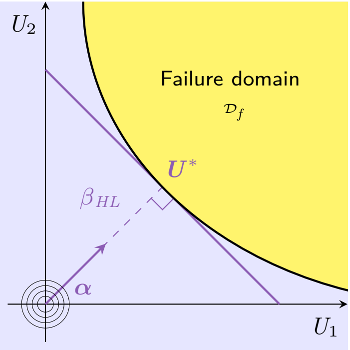
Measure of a subdomain





When measuring a subset (*e.g.* the failure domain) of the Gaussian space, the points that contribute the most to the result are those that are close to the origin

Step 2: Search of the design point



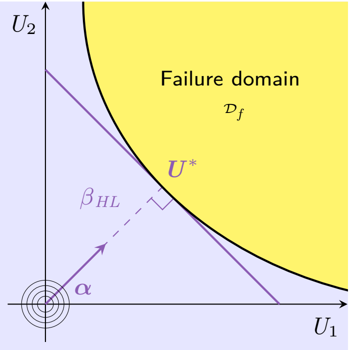
- The design point U^* is defined as the point of the failure domain that is the closest to the origin in the standard normal space.
- It is obtained by solving the constrained optimization problem:

$$U^* = \arg \min_{U \in \mathbb{R}^M} \{\| U \|^2 , \ g(\mathcal{T}(U)) \leq 0\}$$

The design point is the most probable failure point in the standard normal space

- The distance $\beta_{HL} = \| U^* \|$ is the Hasofer-Lind reliability index.
- The unit vector α is defined so that $U^* = \beta_{HL} \, \alpha$

Step 2: Search of the design point



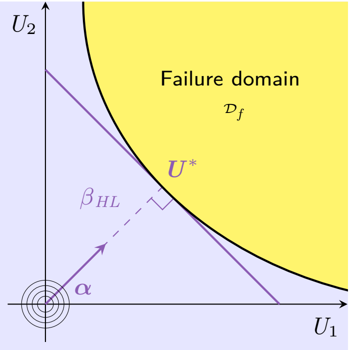
- The design point $\boldsymbol{U}^*$ is defined as the point of the failure domain that is the closest to the origin in the standard normal space.
- It is obtained by solving the constrained optimization problem:

$$\boldsymbol{U}^* = \arg \min_{\boldsymbol{U} \in \mathbb{R}^M} \{\parallel \boldsymbol{U} \parallel^2, \, g(\mathcal{T}(\boldsymbol{U})) \leq 0\}$$

The design point is the most probable failure point
in the standard normal space

- The distance $\beta_{HL} = \parallel \boldsymbol{U}^* \parallel$ is the Hasofer-Lind reliability index.
- The unit vector $\boldsymbol{\alpha}$ is defined so that $\boldsymbol{U}^* = \beta_{HL}\, \boldsymbol{\alpha}$

Step 2: Search of the design point



- The design point U^* is defined as the point of the failure domain that is the closest to the origin in the standard normal space.
- It is obtained by solving the constrained optimization problem:

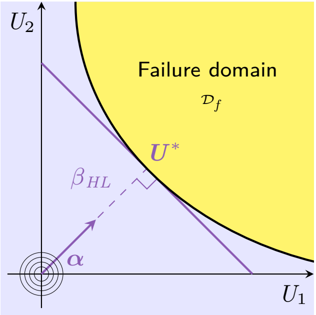
$$U^* = \arg \min_{U \in \mathbb{R}^M} \{ \| U \|^2 , \, g(\mathcal{T}(U)) \leq 0 \}$$

The design point is the most probable failure point in the standard normal space

- The distance $\beta_{HL} = \| U^* \|$ is the Hasofer-Lind reliability index.
- The unit vector α is defined so that $U^* = \beta_{HL} \, \alpha$

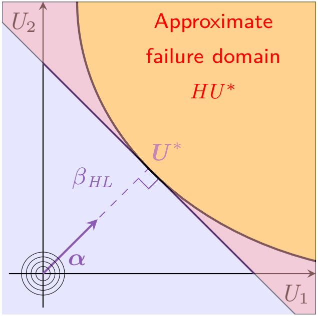
Step 3: FORM approximation

Linearization at the design point



$$P_f = \int_{D_f = \left\{ \boldsymbol{u} \in \mathbb{R}^M \,:\, g(\mathcal{T}(\boldsymbol{u})) \leq 0 \right\}} \varphi_M(\boldsymbol{u}) \, d\boldsymbol{u}$$

Step 3: FORM approximation
Linearization at the design point

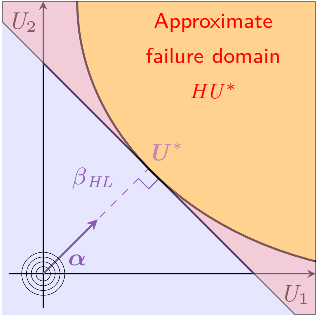


$$P_f = \int_{D_f = \left\{ \boldsymbol{u} \in \mathbb{R}^M \,:\, g(\mathcal{T}(\boldsymbol{u})) \leq 0 \right\}} \varphi_M(\boldsymbol{u}) \; d\boldsymbol{u}$$

The failure domain $\mathcal{D}_f$ is replaced by the half-space that is tangent at the design point $\boldsymbol{U}^*$:

$$P_f \approx \int_{HU^*} \varphi_M(\boldsymbol{u}) \; d\boldsymbol{u}$$

Step 3: FORM approximation
Linearization at the design point



$$P_f = \int_{D_f = \left\{ \boldsymbol{u} \in \mathbb{R}^M \,:\, g(\mathcal{T}(\boldsymbol{u})) \leq 0 \right\}} \varphi_M(\boldsymbol{u}) \; d\boldsymbol{u}$$

The failure domain $\mathcal{D}_f$ is replaced by the half-space that is tangent at the design point $\boldsymbol{U}^*$:

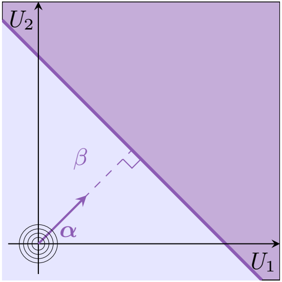
$$P_f \approx \int_{HU^*} \varphi_M(\boldsymbol{u}) \; d\boldsymbol{u}$$

- The halfspace HU^* may be defined by its distance to the origin which is the Hasofer-Lind reliability index β_{HL} and a unit normal vector.

$$HU^* \; : \quad \beta_{HL} - \boldsymbol{\alpha} \cdot \boldsymbol{u} \leq 0$$

- The approximation of the probability of failure reduces to computing the measure of a half-space.

Step 3: FORM approximation
Measure of a half-space



A half-space may be defined by an hyperplane whose reduced equation reads:

$$\mathcal{H}(\boldsymbol{\alpha}, \beta) : \qquad \beta - \boldsymbol{\alpha} \cdot \boldsymbol{u} \leq 0$$

where β is the Euclidean distance of the hyperplane to the origin and $\boldsymbol{\alpha}$ is a unit normal vector.

The (Gaussian) measure of this half-space is:

$$\mathbb{P}\left(\beta - \boldsymbol{\alpha} \cdot \boldsymbol{U} \leq 0\right) = \Phi(-\beta)$$

where Φ is the standard normal CDF: $\Phi(x) = \int_{-\infty}^{x} e^{-t^2/2}/\sqrt{2\pi}\, dt$

FORM in a nutshell

Ingredients

- an iso-probabilistic transform of the input random vector $\boldsymbol{X}$ into a standard normal vector $\boldsymbol{U}$
- the search for the design point $\boldsymbol{U}^*$ in this space (which requires *e.g.* $5 - 10(M+1)$ calls to g)
- the linearization of the limit state surface at the design point and the computation of the approximated failure probability:

$$P_{f,\text{FORM}} = \Phi\left(-\beta_{HL}\right) \qquad \beta_{HL} = \parallel \boldsymbol{U}^* \parallel$$

where β_{HL} is the Hasofer-Lind reliability index.

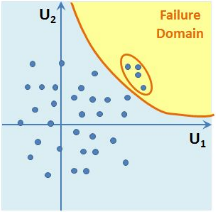
Limitations

- FORM relies upon the unicity of the design point.
- The optimization algorithm may not converge.
- The linear approximation of the limit state surface may be poor.

Outline

Back to Monte Carlo simulation

- Monte Carlo simulation is inefficient for computing small probabilities of failure due to the fact that most sample points are drawn in the vicinity of $\mu_{\boldsymbol{X}}$ whereas failure is related to extreme realizations of $\boldsymbol{X}$.



- After transforming the problem in the standard normal space the probability of failure reads:

$$P_f = \int_{D_f = \left\{\boldsymbol{u} \in \mathbb{R}^M \,:\, g(\mathcal{T}(\boldsymbol{u})) \leq 0\right\}} \varphi_M(\boldsymbol{u}) \, d\boldsymbol{u}$$

- Efficiency may be gained by modifying the sampling scheme in order to concentrate the realizations in the region of interest

Importance sampling

Importance sampling

Principle

- Consider a distribution function $h : \mathbb{R}^M \mapsto \mathbb{R}$ such that $h(\boldsymbol{x}) \neq 0 \; \forall \, \boldsymbol{x} \in \mathcal{D}_f$. Then:

$$\begin{aligned} P_f &= \int_{\mathbb{R}^M} 1_{\mathcal{D}_f}(\boldsymbol{u}) \, \varphi_M(\boldsymbol{u}) \, d\boldsymbol{u} \\ &= \int_{\mathbb{R}^M} \frac{1_{\mathcal{D}_f}(\boldsymbol{u}) \, \varphi_M(\boldsymbol{u})}{h(\boldsymbol{u})} \, h(\boldsymbol{u}) \, d\boldsymbol{u} \\ &= \mathbb{E}_h \left[\frac{1_{\mathcal{D}_f}(\boldsymbol{Z}) \, \varphi_M(\boldsymbol{Z})}{h(\boldsymbol{Z})} \right] \qquad \boldsymbol{Z} \sim h(\boldsymbol{x}) \end{aligned}$$

- h is called the importance sampling or instrumental density.

- It is freely selected provided it is non zero over the failure domain.

Importance sampling estimator

Monte Carlo estimator

$$\hat{P}_{f,IS} = \frac{1}{N} \sum_{i=1}^{N} \frac{1_{\mathcal{D}_f}(\boldsymbol{Z}_i)\, \varphi_M(\boldsymbol{Z}_i)}{h(\boldsymbol{Z}_i)} \qquad \boldsymbol{Z}_i \sim h(\boldsymbol{x}), \text{ i.i.d}$$

- $\hat{P}_{f,IS}$ is unbiased and convergent:

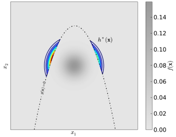
$$\text{Var}\left[\hat{P}_{f,IS}\right] = \frac{1}{N} \text{Var}_h \left[\frac{1_{\mathcal{D}_f}(\boldsymbol{Z})\, \varphi_M(\boldsymbol{Z})}{h(\boldsymbol{Z})}\right]$$

Optimal instrumental density

(Rubinstein, 2008)

- The optimal instrumental density h^* allows one to achieve the minimal variance for $\hat{P}_{f,IS}$:

$$h^*(\boldsymbol{x}) = \frac{\mathbf{1}_{\mathcal{D}_f}(\boldsymbol{x})\, \varphi_M(\boldsymbol{x})}{P_f}$$



The optimal importance sampling density depends on the unknown quantity P_f!

FORM-based importance sampling

Melchers (1989)

- Following the development of FORM, engineers tried to take advantage from the information brought by a FORM analysis in order to build a suitable importance sampling density h.

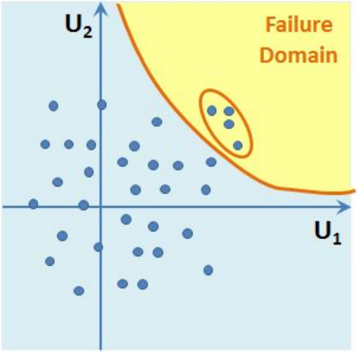
- The design-point importance sampling is based on:
  - the computation of the design point by FORM
  - the use of a shifted multinormal PDF that is centered on $\boldsymbol{U}^*$ as an instrumental density:

$$h(\boldsymbol{x}) = \varphi_M(\boldsymbol{x} - \boldsymbol{U}^*) = (2\pi)^{-M/2} e^{-\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{U}^*\|}$$

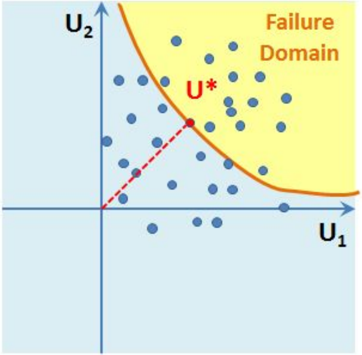
- The IS estimator reads:

$$\begin{aligned}\hat{P}_{f,IS} &= \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{D_f}(\boldsymbol{U}_i) \frac{\varphi_M(\boldsymbol{U}_i)}{\varphi_M(\boldsymbol{U}_i - \boldsymbol{U}^*)} \\ &= \frac{1}{N} e^{-\beta^2/2} \sum_{i=1}^{N} \mathbf{1}_{D_f}(\boldsymbol{U}_i) \exp(-\boldsymbol{U}_i \cdot \boldsymbol{U}^*)\end{aligned}$$

Illustration



Crude Monte Carlo simulation



Design point importance sampling

Conclusion

- Monte Carlo simulation is usually not applicable directly in structural reliability problems due to its computational cost.
- In contrast FORM (and its second-order extension SORM) are very efficient. However no error estimate is available.
- Importance sampling (IS) tries to combine both approaches, *i.e.* it is a simulation method which concentrates the samples in the region of interest.
  - FORM-based IS makes use of a multinormal instrumental density centered on FORM's design point.
  - Other approaches exist, *e.g.* the cross-entropy method.
- Alternative simulation methods such as directional simulation and subset simulation (splitting) have been proposed in the last decade. They remain costly.

In order to compute rare event probabilities using $\approx 100 - 1000$ runs of the limit state function, meta-models are required

Outline

1. Introduction
2. Classical computational methods
3. Metamodels in rare event simulation
    - Kriging
    - Adaptive kriging for structural reliability
    - Meta-model- based importance sampling
4. Application examples

What is a meta-model?

Definition

- A meta-model $\tilde{g}$ is a fast-to-evaluate function that mimics the behaviour of the initial limit state function g, *i.e.* $g(\boldsymbol{x}) \approx \tilde{g}(\boldsymbol{x}) \quad \forall\, \boldsymbol{x} \in A \subset \mathbb{R}^M$.

- It is built using a set of runs of the true limit state function on a so-called experimental design:

$$\mathcal{X} = \left\{ \boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(N)} \right\}$$

  i.e. :

$$\boldsymbol{\Gamma} = \left\{ g\left(\boldsymbol{x}^{(1)}\right), \dots, g\left(\boldsymbol{x}^{(N)}\right) \right\}^{\mathsf{T}}$$

- Experimental designs may be fixed (*e.g.* Latin Hypercube sampling, low-discrepancy sequences, etc.) or adaptively enriched.

Types of meta-models in structural reliability

(Sudret, 2012)

- Polynomial expansions:
  - FORM may be considered as a linear approximation of the limit state function in the standard normal space: $\tilde{G}(\boldsymbol{u}) \approx \beta_{HL} - \boldsymbol{\alpha} \cdot \boldsymbol{U}$.
  - SORM is based on a parabolic second-order expansion.
  - More generally polynomial chaos expansions may be used:

$$\tilde{G}(\boldsymbol{u}) = \sum_{j \in \mathcal{J}} a_j \, \Psi_j(\boldsymbol{u}) \qquad \text{(Orthogonal polynomials)}$$

- Support vector machines: $\tilde{G}(\boldsymbol{u}) = \sum_{j} a_j \, K(\boldsymbol{u}, \boldsymbol{u}_j)$
- Kriging

Kriging surrogate (a.k.a Gaussian process modelling)

Heuristics

Sacks *et al.* , (1989)

The limit state function $y = g(\boldsymbol{x})$ as a function is assumed to be a particular realization of a Gaussian process $Y(\boldsymbol{x}, \omega)$:

$$Y(\boldsymbol{x}, \omega) = \boldsymbol{f}(\boldsymbol{x})^{\mathsf{T}} \boldsymbol{a} + Z(\boldsymbol{x}, \omega)$$

where:

- the mean value is parameterized by a set of prescribed functions $\{f_i, \ i = 1, \ldots, P\}$ (regression part)
- $Z(\boldsymbol{x}, \omega)$ is a zero-mean stationary Gaussian process with variance σ_Y^2 and assumed covariance function:

$$C_{YY}\left(\boldsymbol{x}, \, \boldsymbol{x}'\right) = \sigma_Y^2 \, R\left(\boldsymbol{x} - \boldsymbol{x}', \, \boldsymbol{\theta}\right) \quad \text{e.g.} \quad \sigma_Y^2 \exp\left(\sum_{k=1}^{M} - \left(\frac{x_k - x_k'}{\theta_k}\right)^2\right)$$

The Gaussian measure artificially introduced on $Y(\boldsymbol{x})$ is different from the aleatory uncertainty on the model parameters $\boldsymbol{X}$

Best linear unbiased estimator (BLUE)

Problem statement

- The available data $\mathcal{X} = \left\{ \left(\boldsymbol{x}^{(i)}, y^{(i)} = g\left(\boldsymbol{x}^{(i)}\right)\right),\ i = 1, \dots, N \right\}$ is a set of pointwise observations of the specific trajectory $g(\boldsymbol{x}) = Y(\boldsymbol{x}, \omega_0)$.
- In other words, $\boldsymbol{\Gamma} = \left\{ g\left(\boldsymbol{x}^{(1)}\right), \dots, g\left(\boldsymbol{x}^{(N)}\right) \right\}^{\mathsf{T}}$ is a realisation of a Gaussian vector $\mathfrak{Y} = \{ Y_1, \dots, Y_N \}$ where $Y_i \equiv Y(\boldsymbol{x}_i, \omega)$.
- Of interest is the prediction of $Y_0 \equiv Y(\boldsymbol{x}, \omega)$ for other points $x \in \mathcal{D}_{\boldsymbol{X}}$.
- The BLUE is cast as:

$$\hat{Y}_0 = \sum_{i=1}^{M} a_i(\boldsymbol{x})\ Y_i$$

such that it is unbiased : $\mathbb{E}\left[\hat{Y}_0 - Y_0\right] = 0$ with minimum variance $\mathbb{E}\left[\left(Y_0 - \hat{Y}_0\right)^2\right]$

Kriging surrogate
Solution

Mean predictor

$$\tilde{g}(\boldsymbol{x}) \stackrel{\text{def}}{=} \mu_{\widehat{Y}}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^{\mathsf{T}}\ \hat{\boldsymbol{a}} + \boldsymbol{r}(\boldsymbol{x})^{\mathsf{T}}\mathbf{R}^{-1}(\boldsymbol{\Gamma} - \mathbf{F}\,\hat{\boldsymbol{a}})$$

where:

$$\begin{aligned} r_i(\boldsymbol{x}) &= R\left(\boldsymbol{x} - \boldsymbol{x}^{(i)}, \boldsymbol{\theta}\right),\ i = 1, \ldots, N \\ \mathbf{R}_{ij} &= R\left(\boldsymbol{x}^{(i)} - \boldsymbol{x}^{(j)}, \boldsymbol{\theta}\right),\ i = 1, \ldots, N,\ j = 1, \ldots, N \\ \mathbf{F}_{ij} &= f_j\left(\boldsymbol{x}^{(i)}\right),\ i = 1, \ldots, p,\ j = 1, \ldots, N \end{aligned}$$

The result is independent of the choice of the properties of the Gaussian process, *i.e.* whatever $\boldsymbol{a}$, σ_Y^2, $\boldsymbol{\theta}$

Kriging variance

$$\sigma^2_{\widehat{Y}}(\boldsymbol{x}) = \sigma_Y^2\left(1 - \left\langle\ \boldsymbol{f}(\boldsymbol{x})^{\mathsf{T}}\quad \boldsymbol{r}(\boldsymbol{x})^{\mathsf{T}}\ \right\rangle \begin{bmatrix} \mathbf{0} & \mathbf{F}^{\mathsf{T}} \\ \mathbf{F} & \mathbf{R} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{f}(\boldsymbol{x}) \\ \boldsymbol{r}(\boldsymbol{x}) \end{bmatrix}\right)$$

Estimation of the parameters

(Santner *et al.* , 2003)

Unknown parameters

$\boldsymbol{a}$: coefficients of the regression part

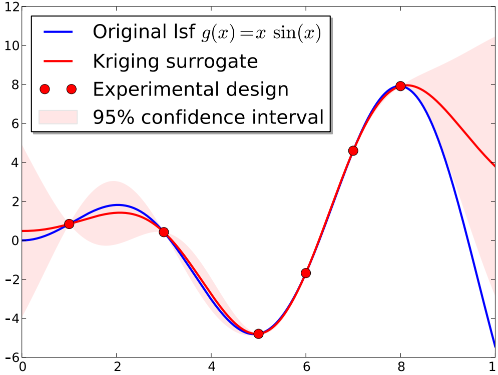
σ_Y^2: variance of the process

$\boldsymbol{\theta}$: correlation lengths in the covariance function

Maximum likelihood estimation

- The likelihood function is obtained from the joint Gaussian distribution of $\{Y_1, \ldots, Y_N\}$.
- A single realization is available, namely the vector of observations $\boldsymbol{\Gamma} = \left\{g\left(\boldsymbol{x}^{(1)}\right), \ldots, g\left(\boldsymbol{x}^{(N)}\right)\right\}^{\mathsf{T}}$.
- Analytical solutions are available for $\widehat{\boldsymbol{a}}$ and $\sigma_{\widehat{Y}}^2$ conditionally to $\boldsymbol{\theta}$. The maximization w.r.t $\boldsymbol{\theta}$ is carried out numerically.

Visualization of a kriging surrogate



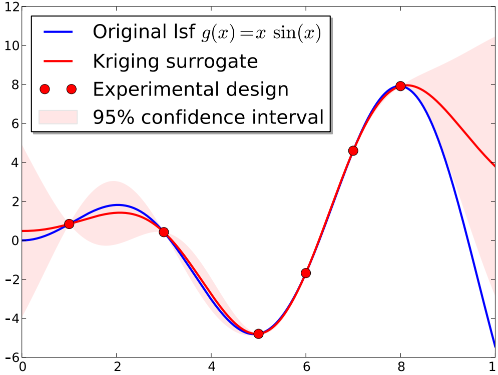
- The surrogate $\mu_{\widehat{Y}}$ interpolates the function g on the experimental design:

$$\mu_{\widehat{Y}}\left(\boldsymbol{x}^{(i)}\right) = g\left(\boldsymbol{x}^{(i)}\right)$$
$$\sigma^2_{\widehat{Y}}\left(\boldsymbol{x}^{(i)}\right) = 0$$

- Due to gaussianity confidence intervals may be drawn.

Kriging provides a built-in estimation
of the (epistemic) error of the surrogate

Visualization of a kriging surrogate



- The surrogate $\mu_{\widehat{Y}}$ interpolates the function g on the experimental design:

$$\mu_{\widehat{Y}}\left(\boldsymbol{x}^{(i)}\right) = g\left(\boldsymbol{x}^{(i)}\right)$$
$$\sigma^2_{\widehat{Y}}\left(\boldsymbol{x}^{(i)}\right) = 0$$

- Due to gaussianity confidence intervals may be drawn.

Kriging provides a built-in estimation
of the (epistemic) error of the surrogate

Outline

1. Introduction
2. Classical computational methods
3. Metamodels in rare event simulation
   - Kriging
   - Adaptive kriging for structural reliability
   - Meta-model- based importance sampling
4. Application examples

Kriging surrogate and active learning

- The kriging variance yields an estimation of the accuracy of the meta-model which may be used in an active learning context.
- The experimental design is enriched iteratively in regions which are meaningful for evaluating the probability of failure, *i.e.* the vicinity of the limit state surface $g(\boldsymbol{x}) = 0$.

Enrichment criteria

- expected feasibility function (Bichon *et al.* (2008) ; Bect *et al.* (2011))

$$EF(\boldsymbol{x}) = \mathbb{E}\left[Feas(\boldsymbol{x})\right] \qquad Feas(\boldsymbol{x}) = \max\left\{\varepsilon - |\widehat{Y}(\boldsymbol{x})|\,,\, 0\right\}$$

- Learning function (Echard *et al.* , 2011-12)

$$U(\boldsymbol{x}) = \frac{|\mu_{\widehat{Y}}(\boldsymbol{x})|}{\sigma_{\widehat{Y}}(\boldsymbol{x})}$$

- Probabilistic classification function (Dubourg *et al.* , 2011-12)

Probabilistic classification function

Definition

$$\pi(\boldsymbol{x}) = \mathcal{P}\left[\hat{Y}(\boldsymbol{x}) \leq 0\right] = \Phi\left(\frac{0 - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right)$$

$\mathcal{P}$ is the Gaussian measure associated with the Gaussian process

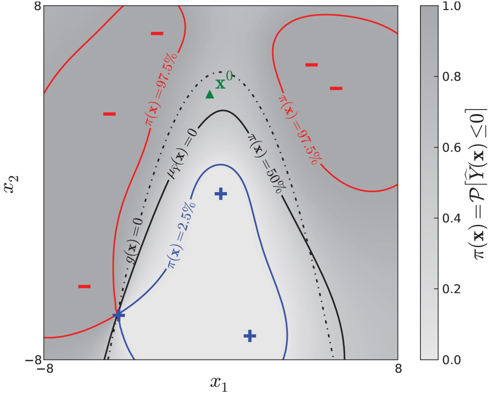
Interpretation

Assume the surrogate is "good" for a specific $\boldsymbol{x}_0$ ($\sigma_{\widehat{Y}}(\boldsymbol{x}_0) \to 0^+$):

- If $\mu_{\widehat{Y}}(\boldsymbol{x}_0) \approx g(\boldsymbol{x}_0) > 0$ then $\pi(\boldsymbol{x}_0) \approx 0$
- If $\mu_{\widehat{Y}}(\boldsymbol{x}_0) \approx g(\boldsymbol{x}_0) < 0$ then $\pi(\boldsymbol{x}_0) \approx 1$

$\pi(\boldsymbol{x})$ is a smooth surrogate of the indicator function $\mathbf{1}_{\mathcal{D}_f}(\boldsymbol{x})$

Margin of uncertainty on the limit state surface



The margin of uncertainty $\mathfrak{M}$ is defined by the $(1-\alpha)$-confidence region of the surrogate limit state surface $\mu_{\widehat{Y}} = 0$, *i.e.* the set of points such that:

$$\alpha/2 \leq \pi(\boldsymbol{x}) \leq 1 - \alpha/2$$

$$\mathfrak{M} = \left\{ \boldsymbol{x} \,:\, -k\,\sigma_{\widehat{Y}}(\boldsymbol{x}) \leq \mu_{\widehat{Y}}(\boldsymbol{x}) \leq +k\,\sigma_{\widehat{Y}}(\boldsymbol{x}) \right\}, \quad k = \Phi^{-1}(1-\alpha/2) \quad \text{e.g. } 1.96$$

Enrichment in the margin of uncertainty

The enrichment criterion $\mathcal{C}(\boldsymbol{x})$ is defined as the (Gaussian) measure of the margin in each point $\boldsymbol{x}$.

$$\mathcal{C}(\boldsymbol{x}) = \mathcal{P}\left[-k\,\sigma_{\widehat{Y}}(\boldsymbol{x}) \leq \widehat{Y}(\boldsymbol{x}) \leq k\,\sigma_{\widehat{Y}}(\boldsymbol{x})\right]$$

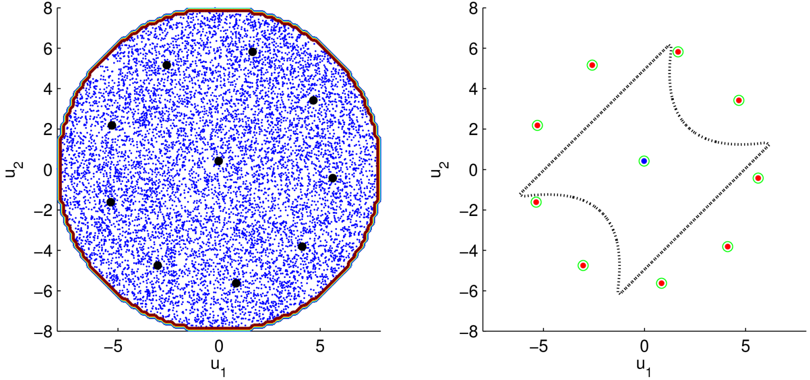
- It could be maximized in order to find the next point to add to the current experimental design.
- It may better be used as a (improper) sampling density in order to draw candidate points for the enrichment (Markov chain Monte Carlo simulation):

$$f_{\mathcal{C}}(\boldsymbol{x}) \propto \mathcal{C}(\boldsymbol{x})\, f_X(\boldsymbol{x})$$

- A batch of reduced size is obtained by K-means clustering.

Sampling in the margin

$$\mathcal{C}(\boldsymbol{u}) = \mathcal{P}[\boldsymbol{u} \in \mathbb{M}] \, \mathbb{1}_{\sqrt{\boldsymbol{u}^\mathsf{T} \boldsymbol{u}} \leq \beta_0}(\boldsymbol{u})$$



Estimators of P_f by substitution

Classical approach

- At each step of the active learning, the probability of failure may by estimated by substituting for the Kriging surrogate $\tilde{g} \equiv \mu_{\widehat{Y}}$ into the definition of the probability of failure:

$$P_f \approx \tilde{P}_f = \mathbb{P}\left(\tilde{g}(\boldsymbol{X}) \leq 0\right) = \int_{\tilde{D}_f = \left\{\boldsymbol{x}: \quad \mu_{\widehat{Y}}(\boldsymbol{X}) \leq 0\right\}} f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x}$$

- Monte Carlo simulation may be used now since evaluating the surrogate $\mu_{\widehat{Y}}(\boldsymbol{x})$ is inexpensive.

- Bounds denoted by $\tilde{P}_f^- / \tilde{P}_f^+$ may also be computed by using $\mu_{\widehat{Y}}(\boldsymbol{x}) \pm k\, \sigma_{\widehat{Y}}(\boldsymbol{x})$ as a surrogate.

Meta-IS: the kriging surrogate is used as a tool for deriving a quasi-optimal importance sampling density.

Outline

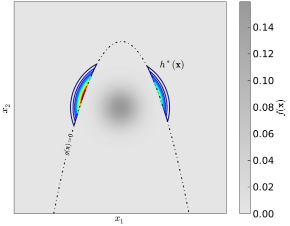
Reminder on importance sampling

Definition

$$P_f = \int_{\mathbb{R}^M} \mathbf{1}_{\mathcal{D}_f}(\boldsymbol{x}) \, \frac{f_X(\boldsymbol{x})}{h(\boldsymbol{x})} h(\boldsymbol{x}) \, d\boldsymbol{x} = \mathbb{E}_h \left[\mathbf{1}_{\mathcal{D}_f}(\boldsymbol{X}) \frac{f_X(\boldsymbol{X})}{h(\boldsymbol{X})} \right]$$

The optimal IS density reads: Rubinstein (2008)

$$h^*(\boldsymbol{x}) = \frac{\mathbf{1}_{\mathcal{D}_f}(\boldsymbol{x}) \, f_X(\boldsymbol{x})}{P_f}$$



$g(x_1, x_2) = 5 - x_2 - \frac{1}{2}(x_1 - 0.1)^2$

- It is not tractable in practice since it involves the unknown P_f!
- It may be approximated using the kriging surrogate.

Quasi-optimal IS density

Proposed IS density:

(Dubourg *et al.* , 2012)

$$h^*(\boldsymbol{x}) = \frac{\mathbf{1}_{\mathcal{D}_f}(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})}{P_f} \quad \rightsquigarrow \quad \tilde{h}(\boldsymbol{x}) \equiv \frac{\pi(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})}{P_{f\varepsilon}} \qquad \pi(\boldsymbol{x}) = \Phi\left(\frac{-\mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right)$$

where the augmented probability of failure $P_{f\varepsilon}$ reads:

$$P_{f\varepsilon} = \mathbb{E}\left[\pi(\boldsymbol{X})\right] = \int_{\mathbb{R}^M} \pi(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x}$$

Unbiased estimator of P_f:

$$P_f = \int_{\mathbb{R}^M} \mathbf{1}_{\mathcal{D}_f}(\boldsymbol{x})\, \frac{f_{\boldsymbol{X}}(\boldsymbol{x})}{\tilde{h}(\boldsymbol{x})}\, \tilde{h}(\boldsymbol{x})\, d\boldsymbol{x} = P_{f\varepsilon} \cdot \underbrace{\int_{\mathbb{R}^M} \frac{\mathbf{1}_{\mathcal{D}_f}(\boldsymbol{x})}{\pi(\boldsymbol{x})}\, \tilde{h}(\boldsymbol{x})\, d\boldsymbol{x}}_{\alpha_{corr}}$$

$$P_f = P_{f\varepsilon} \cdot \mathbb{E}_{\tilde{h}}\left[\frac{\mathbf{1}_{\mathcal{D}_f}(\boldsymbol{x})}{\pi(\boldsymbol{x})}\right]$$

Monte Carlo estimator

The meta-IS estimator of P_f is the product of two terms, namely the augmented probability of failure and a correction factor:

$$\widehat{P_f} = \widehat{P_{f\varepsilon}} \cdot \widehat{\alpha}_{corr}$$

$$\widehat{P_{f\varepsilon}} = \frac{1}{N_\varepsilon} \sum_{l=1}^{N_\varepsilon} \pi(\boldsymbol{x}^{(l)})$$

- computed from the kriging surrogate (inexpensive if $N_\varepsilon \sim 10^{3-4}$)
- $\boldsymbol{x}^{(l)} \sim f_{\boldsymbol{X}}(\boldsymbol{x})$

$$\widehat{\alpha}_{corr} = \frac{1}{N_{corr}} \sum_{k=1}^{N_{corr}} \frac{\mathbf{1}_{\mathcal{D}_f}(\tilde{\boldsymbol{x}}^{(k)})}{\pi(\tilde{\boldsymbol{x}}^{(k)})}$$

- computed from the original "true" limit state function
- $\boldsymbol{x}^{(k)} \sim \tilde{h}(\boldsymbol{x})$

Interpretation

The correction factor emphasizes the samples that are misclassified by the smoothed kriging-based limit state function π.

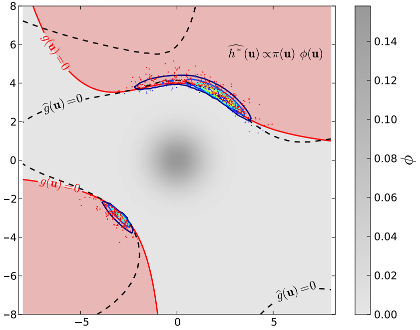
Two-dimensional series system

Au & Beck (1999)

Limit state function

$$g(x_1, x_2) = \min \left\{ c - 1 - x_2 + e^{-x_1^2/10} + \left(\frac{x_1}{5}\right)^4 \; , \; \frac{c^2}{2} - x_1\, x_2 \right\}$$

where $X_1, X_2 \sim \mathcal{N}(0, 1)$.



Three case studies :
$c = 3, 4$ or 5

Two-dimensional series system

Results

Method		Monte Carlo (ref)	Subset	Meta-IS[1]
	N	10^7	300,000	44 + 600
$c = 3$	p_f	3.48×10^{-3}	3.48×10^{-3}	3.54×10^{-3}
	C.o.V.	0.5%	<3%	<5%
	N	10^8	500,000	64 + 600
$c = 4$	p_f	8.94×10^{-5}	8.34×10^{-5}	8.60×10^{-5}
	C.o.V.	3.3%	<4%	<5%
	N	10^9	700,000	40 + 2,900
$c = 5$	p_f	9.28×10^{-7}	6.55×10^{-7}	9.17×10^{-7}
	C.o.V.	3.3%	<5%	<5%

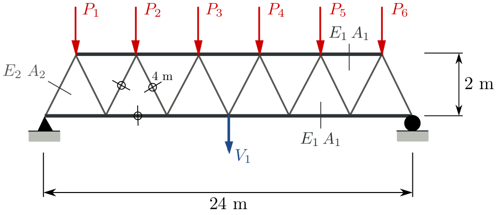
About 3% accuracy on P_f (less than 0.2% error on β) in the range $[10^{-7}, 10^{-3}]$

[1] $N_{tot} = N + N_{\text{IS}}$.

Finite element reliability analysis

Truss structure

Blatman (2009)



$$g(\boldsymbol{X}) = V_1 - FEM(\boldsymbol{X})$$

$$\boldsymbol{X} = \{E_1, E_2, A_1, A_2, P_1, ..., P_6\}^{\mathsf{T}}$$

Variable	Distribution	Mean	C.V
E_1, E_2 (Pa)	Lognormal	2.10×10^{11}	10%
A_1 (m^2)	Lognormal	2.0×10^{-3}	10%
A_2 (m^2)	Lognormal	1.0×10^{-3}	10%
P_1-P_6	Gumbel	5.0×10^{4}	15%

Finite element reliability analysis

Results

Threshold (cm)		Importance sampling	FORM	Meta-IS[a]
		Blatman (2009)		
10	N_{tot}	500,000	121	160 +31
	P_f	4.00×10^{-2}	2.81×10^{-2}	4.35×10^{-2} (C.o.V.=1.2%)
	β	1.75	1.91	1.71
14	N_{tot}	500,000	121	160 +31
	P_f	3.45×10^{-5}	1.28×10^{-5}	3.47×10^{-5} (C.o.V.=3.7%)
	β	3.98	4.21	3.98

[a] $N_{tot} = N + N_{\text{IS}}$.

- About the same cost as FORM
- Unbiased estimation of P_f within 1% accuracy (on P_f!!)

Summary

- The quantification of rare events probabilities is of great importance in civil & mechanical engineering since it is related to the reliability of the systems under consideration.
- The probability of failure is cast as a multidimensional integral whose direct computation is not possible due to the implicit definition of the failure domain.
- Crude Monte Carlo simulation is not efficient and in practice not applicable due to unaffordable computational costs.
- Advanced simulation methods based on importance sampling and subset simulation are still too expansive in many situations. The only solution is then to use surrogate models.

Summary

- Kriging (a.k.a Gaussian process modelling) is a type of surrogate models that provides an error indicator which may be used in the context of active learning (adaptive experimental designs).
- The Kriging variance is used for two purposes:
    - define a probabilistic classification function $\pi(\boldsymbol{x}) = \Phi\left(\dfrac{0 - \mu_{\widehat{Y}}(\boldsymbol{x})}{\sigma_{\widehat{Y}}(\boldsymbol{x})}\right)$ which is used in order to enrich the experimental design.
    - define "confidence intervals" on the surrogate models, *e.g.* $\mu_{\widehat{Y}}(\boldsymbol{x}) \pm k\,\sigma_{\widehat{Y}}(\boldsymbol{x})$ which allows one to compute (not necessarily strict) bounds on P_f.
- In most current approaches there is no proof that the probability of failure computed by substituting $\mu_{\widehat{Y}}$ for g is unbiased.

Summary

- In meta-model-based importance sampling (meta-IS), Kriging is used as a tool for deriving a quasi-optimal importance sampling density.
- An unbiased estimator of P_f is obtained as the product of the augmented probability of failure $P_{f\varepsilon} = \mathbb{E}_{\boldsymbol{X}}\left[\pi(\boldsymbol{X})\right] = \int_{\mathbb{R}^M} \pi(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, d\boldsymbol{x}$ and a correction factor.
- Although $P_{f\varepsilon}$ is often a good estimation of P_f, the correction factor ensures that the estimator is unbiased by accounting for the possible misclassification of certain points by the surrogate limit state function.

Thank you very much for your attention !